1. Introduction

This work presents a series of methods for the automatic identification of designers’ activities through the monitoring of the digital tools that they use. Data extracted by monitoring the execution of commands and the flow of information between user and computer can be used to form a machine-readable profile of design sessions. The profiles can be used for the purpose of providing additional computer-based support to the user during a design session or for the storage and reuse of information where reference to historic information such as the understanding of previous design rationale is required. Initially, we explore the implications of profiling design activities, particularly with respect to the potential impact on the lifecycle of design information. A methodology for profiling design activities is then presented, along with some initial development work directed towards its implementation and use.

2. Objectives

The objectives of this research are, in more detail:

- To investigate information that can be extracted from aspects of computer-based design sessions that may provide useful and useable information about the design task being carried out. More specifically, the work involves the evaluation of all elements of the human-computer interface with respect to their usefulness in (i) determining the current goals of a designer and (ii) providing a description of events that can be stored and reused for reference in similar future design scenarios.

- To establish methods for recording transactions in computer-based applications that are frequently used in the embodiment and detail phases of the design process. The applications considered include computer-aided engineering software and the tools necessary for the access of local, networked and Internet information repositories. The methods being explored include: (i) the recording of the order and timing of access to information used in decision-making processes, in parallel with the creation of formal design representations, (in CAD Computer-Aided Design for example) (ii) establishing a user’s current focus or goal through an interpretation of actions observed at the human-computer interface (e.g. keyboard input and changes to information displayed on screen).
It is intended that the methods established are used to form design activity profiles that have the following purposes:

1. To enable the “push”\(^1\) of information contextually relevant to the user’s current task by matching a computer generated description of the current design situation with previous design situations and associated information retrieved from similar historical contexts.

2. To capture and store a description of the current activity. This historical record of information can be used to “push” information to future users whose current activity resembles a similar context.

3. To record a log of actions carried out in design activities that can be used to help trace rationales for decisions taken in the past.

The primary contribution in this paper is the presentation of a methodology that can be used to create and store design activity profiles. The information push aspects are also discussed, but are considered in full detail in a related publication [1].

Before discussing how the creation of design activity profiles might be achieved, the following section explores the possible positive and negative aspects of profiling design activities through consideration of the lifecycle of design information.

3. The lifecycle of design information

This section is divided into 3 sub-sections: in 3.1 the concept of the lifecycle of information in design is introduced; in 3.2 the impact of the proposed methodology on design information lifecycles is considered; finally, in 3.3 the positive and negative aspects of implementing the methodology are considered.

3.1 Introduction to the lifecycle of design information

In this section we consider the use of information in the design process, in particular exploring the variation of its use in respect of differing degrees of design originality.

A useful starting point is the consideration of a general model of the lifecycle of information. A number of general information lifecycles have been proposed [2][3], however [4] proposes a model particularly suitable for this analysis. Two aspects of this model are of interest: the continuous recycling process shown by the loop (2,4,5) and the continuous addition and deletion of information from this loop shown by (1,2,3).

![Figure 1. General information lifecycle model [4]](image)

---

\(^1\) Information “push” refers to the delivery of information in anticipation of a user’s needs. This is in contrast to a “pull” approach where a user requests and receives a specific piece of information [5]
Pahl and Beitz distinguish between 3 types of design process, original, adaptive and variant design [6]. It is possible to examine the role of the information lifecycle in the context of these types of design (as long as a reductionist or simplified view of information inputs and outputs is assumed).

Figure 2. Design information lifecycle in original design

In the case of original design (see Figure 2), it is proposed that information inputs primarily originate from the ad-hoc collection, creation and capture of information (1). The actual process of design is represented by the organisation of this information (2) facilitated by memory and knowledge of the designers. The information output (4), is typically embodied in computer-aided models of the design and other document types which form a formal, structured information based representation of the design. Finally, all information about and contained within (1) and (2) that is not included in the formal representation may be disposed of or left in an unstructured form that is not readily reused. An obvious example in this respect is information regarding alternative design options considered but rejected in the final design (in the case that this information is omitted in design reports).

Figure 3. Design information lifecycle in adaptive & variant design

In adaptive or variant design, which constitutes the majority of new designs [5], the starting point is an existing design principle, which is adapted or varied to meet an alternate set of requirements. In this case, the information inputs are a combination of the formal representation of the existing design(s) (5), and additional information collected, created, received and captured by the designer. The information output, (4), and information disposed, (3), are the same as in the case of original design.

3.2 Impact of the proposed methodology on the design information lifecycle

Having illustrated the major design types in light of their respective information lifecycles the possible impact of the proposed methodology on this model can be introduced.
This research is concerned with changing 2 aspects of the design information lifecycle; how to reduce the proportion of information that is disposed (4), and how to re-use it more effectively by introducing it when it is timely and relevant to the designers’ activities (6). The objective is to keep as much related information “within the loop”, including the unstructured and informal information that may or may not be repeated in the formal description of the design.

The resulting lifecycle is likely to contain a significantly greater proportion of redundant or repeated information. The next section discusses the implications of this characteristic along with some other general remarks regarding benefits and limitations of the approach.

3.3 Conclusions

The possible drawbacks of retaining more information than is necessary to sustain the resulting design (in manufacture) include the added cost of sustaining larger information archives, and arguments relating to reduced efficiency resulting from “information overload” or becoming lost in too much information [7][8].

However, information that is redundant for sustaining a design (in manufacture) may not be redundant for sustaining the evolution of that design. For example, information about a design feature that was considered, but not included in an original design maybe more suitable in an adapted configuration considered at a later point in time, and so information about the feature is worth retaining.

In conclusion, the costs involved in allowing for redundancy must be compared with possible added value to a design and its evolution resulting from the adoption of such an approach. On consideration of the problems associated with “information overload”, it is thought that the effective organisation and possibly artificial intelligence in the dissemination of information may go some way to alleviating this. These are aspects addressed by the proposed methodology to some extent and are discussed in more detail in later sections.

4. Methodology for profiling computer based design activities

The approach taken has been to systematically assess the characteristics of computer-based design activities to fine levels of granularity to gain an understanding of the potential constituents of design profiles. Initially, four broad categories of activity were identified: CAD, technical writing, electronic correspondence and information searching and browsing. For each category a number of potentially recordable profile elements, defined as data items to be recorded or monitored, were identified. The methodology requires that as the profile elements are captured during the course of a design session they are associated with nodes that are used to describe aspects of the task being carried out by the designer.
In 4.1 and 4.2 the concepts of profile elements and nodes are introduced respectively. In 4.3 some examples of how the resulting design profiles can be used to improve the effectiveness and efficiency of information management in design are considered.

4.1 Profile elements as constituents of design activity profiles

The aim of this section is to introduce the concept of profile elements as being the constituents of design activity profiles. More specifically, the characteristics of computer based actions that qualify them as being useful constituents of the resulting profiles are considered. Some coverage on the technological means for the capture of relevant profile elements is also given.

The profile elements exhibited in computer-based activities refer to all information describing the activity context that can be captured implicitly (i.e. without the user explaining actions) through a software process, and which can then be used or stored without hindrance to the user. Two broad categories of profile element (PE) have been defined.

1. **Information Usage PE**: those elements that describe information usage or change. This covers all uses and changes of information occurring as a result of interaction with the information space where the activity is being carried out. The term “information space” in this discussion is defined in two ways. Conceptually, it is used to refer to the information being observed or manipulated in the task being carried out. In another sense it refers to the usage and manipulation of data contained within files. These elements need not consist of a duplicate of the information but a description of where the information resides.

2. **Interface Action PE**: those elements that describe the interface actions carried out in the current activity. This covers interactions with devices such as the mouse and keyboard.

Table 1. Categorisation of the major profile elements considered in this work

<table>
<thead>
<tr>
<th>PE type</th>
<th>PE description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Usage PE</td>
<td>Files created, renamed or destroyed</td>
</tr>
<tr>
<td></td>
<td>Location of file / web page in focus</td>
</tr>
<tr>
<td></td>
<td>Meta-data of file / document in focus</td>
</tr>
<tr>
<td></td>
<td>Parse-able text from window in focus</td>
</tr>
<tr>
<td></td>
<td>CAD model construction sequence</td>
</tr>
<tr>
<td></td>
<td>Screenshots / images occurring in window in focus</td>
</tr>
<tr>
<td>Interface Action PE</td>
<td>Resizing of windows / applications on interface display</td>
</tr>
<tr>
<td></td>
<td>Window / application focus</td>
</tr>
<tr>
<td></td>
<td>Focus / selections inside applications</td>
</tr>
<tr>
<td></td>
<td>Application start / end</td>
</tr>
<tr>
<td></td>
<td>Menu selections</td>
</tr>
<tr>
<td></td>
<td>Keyboard command shortcuts</td>
</tr>
<tr>
<td></td>
<td>Function keys</td>
</tr>
<tr>
<td></td>
<td>Text entry in dialogue box</td>
</tr>
<tr>
<td></td>
<td>Command text entered at command prompt</td>
</tr>
<tr>
<td></td>
<td>Text written to document or file</td>
</tr>
</tbody>
</table>

Table 1. Categorisation of the major profile elements considered in this work
Table 1 shows a number of general profile elements in these 2 categories. All the extractable profile elements listed provide some form of information about the task being carried out. However, this research is only interested in the elements that describe the purpose of the activity. It is proposed that the profile elements that an activity exhibits include some that are highly contextually relevant to the activity purpose but others that have no relevance at all. For example, in the task of writing up research obtained from the Internet, the opening and closing of applications such as Word™, the resizing of windows, and the execution of back and forward commands in Internet Explorer™ are not contextually relevant to the purpose of the task. However, the search terms and criteria entered or selected in a search engine, the location and content of web pages accessed, and the text entered in applications all are relevant.

The usefulness of profile elements can be related to the 3 general purposes of profiling design activities, which are restated here in the context of profile elements for clarity:

1. To “push” information contextually relevant to the user’s current task by matching profile elements exhibited by the current task with profile elements and associated information retrieved from similar historical contexts.

2. To capture and store any available profile elements that describe the current activity. This historical record of information used can be used to “push” information to future users whose current activity resembles a similar context.

3. To record a log of actions by storing Information Usage PEs. The resulting record can be used to help trace rationales for decisions taken in historical tasks.

Generally speaking the Information Usage PEs are useful for (2) and (3) where a record of information used and manipulated can be stored in a profile of the activity for later reuse. The Interface Action PEs are more useful for (1), particularly forms of text entry which can be used to form queries that represent a partial description of the current activity.

Table 2 lists some of the technologies that enable the capture of profile elements useful in this research.

<table>
<thead>
<tr>
<th>Profile Element</th>
<th>Software Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parse-able onscreen Text (Web based)</td>
<td>IBM Web intermediary (WBI) [9] for documents opened in web based browser</td>
</tr>
<tr>
<td>Document Metadata</td>
<td>IBM Web intermediary (WBI), Microsoft Active Accessibility [10]</td>
</tr>
<tr>
<td>URL Capture</td>
<td>IBM Web intermediary (WBI)</td>
</tr>
<tr>
<td>Local / Networked document access &amp; modification</td>
<td>Microsoft Windows audit policy / .NET framework</td>
</tr>
<tr>
<td>Model Construction Sequence</td>
<td>CAD package APIs</td>
</tr>
<tr>
<td>Feature / Object Creation</td>
<td>CAD package APIs</td>
</tr>
<tr>
<td>Document Implicit Indicator (e.g. doc. browsing time)</td>
<td>Microsoft Active Accessibility</td>
</tr>
<tr>
<td>Typed Text</td>
<td>Microsoft Active Accessibility</td>
</tr>
<tr>
<td>Command Shortcuts</td>
<td>Microsoft Active Accessibility</td>
</tr>
<tr>
<td>Menu Selections</td>
<td>Microsoft Active Accessibility</td>
</tr>
</tbody>
</table>
4.2 Nodes that provide the framework for the organisation of design profiles

This section discusses the way in which profile elements captured from computer based design sessions can be associated with nodes that provide the framework for storing design profiles. There is only a need to store the Information Usage PEs beyond the point at which they are captured as the Interface Action PEs are generally only of use at the time that they occur.

The approach taken in the methodology is to attach profile elements, representing unstructured information in documents used, to node(s) which describe the activity being carried out. 3 types of node have been identified as being useful in computer based activities, nodes that associate information with a file, an abstract concept, or an interest group. A brief description of the each node type is given here:

- **File / document nodes**: this node is associated with the file that is being worked on that typically will contain the formal information representation or result of the activity. So in the case of writing a technical report the node is defined as being connected or associated with the file itself.

- **Abstract concept**: this type of node refers to a something which is recognised by the individual or a group of people as referring to a concept, for example, a project or product name. It does not necessarily refer to any file system being worked on but maybe an area which a team is working in.

- **Interest Group**: this refers to an area of interest subscribed to by a number of people. For example topics of interest to a community of practice. The interest group node refers to a field of interest with greater longevity than an abstract concept node, which, for example, may become redundant at the closure of a project.

Figure 5 shows a schematic representation of a design activity without any node associations. The unstructured information inputs refer to the ad-hoc collection, creation retrieval and capture of information in the design information lifecycle (discussed in 3.1). The formal information outputs refer to the structured information based representation of the design resulting from the activity (discussed in 3.1).

**Figure 5. Schematic representation of design activities without any node associations**

Figure 6 shows the same design activities with the node associations in place. Here the, unstructured information inputs, in the form of captured profile elements are assigned to relevant nodes. Node(s) may be selected from a library of nodes created or used by the individual or project team.
The greatest barrier to implementing a system that enables these node connections is in the design of a system that does so without being of hindrance to the designer in their task. There are several implementation options in this regard, ranging from approaches which are fully automatic but inaccurate, to those which are semi-automatic or manual. The difficulty lies not in the ability to actually capture the data but to decide whether it useful and to which node it should be assigned.

A semi-automatic process that is being considered is to allow the user to set “active node(s)” for periods of time during the design session. During these periods all information usage based PEs captured are associated with the node(s) currently active. An important aspect for this strategy to be successful is to allow the user set and switch between active node(s) and profile elements easily and quickly. Another important aspect is to allow the user to review and modify the information that is to be submitted to a node at the end or during the session in order to make the process completely transparent. Figure 7 shows a possible sequence of actions using this approach.

There are further steps that can be taken to maximize the usefulness of the resulting activity profile through the configuration of subsets of profile elements to be recorded to suit different types of activity. For example when editing a document, profile elements for recently typed text and text parsed from web pages may be useful to capture. In the case of computer aided design profile elements capturing the order and timing of CAD component creation in parallel with the order and timing of information used may be considered to be important.

Where groups of similar types of task are established, predefined configurations can then be developed to suit the activity.
4.3 Use of design activity profiles to push information

The design activity profiles allow for extra information to be preserved in the lifecycle. However, with a potentially much larger information base at the disposal of a designer it becomes more important to provide information that is relevant and timely depending on the current design situation. This section discusses strategies for adding artificial intelligence in the dissemination of information contained within design profiles to the user.

After a period of time it is envisaged that for the nodes that are being actively used by an individual or group, associated information bases will be established. The information bases will consist of electronic references to indexed files and documents developed through the capture of relevant profile elements.

There are 2 possibilities for the exploitation of this information in circumstances where it might be needed. In the first case, an information base could be manually searched or queried by the computer user who has an information need related to the associated node. Alternatively, relevant and timely information could be presented to the computer user autonomously using knowledge of the currently active node(s) and an interpretation of the current information need based on profile elements captured in the design session.

A possible scenario of the autonomous approach is shown schematically in Figure 8. Initially, the profile elements (interface action and information usage) captured in the current task are used, where possible, to form a query intended to describe the current information need of the user (1). A search query for related information is then executed on the information base associated with the active node (2). Lastly, the user is notified or updated of possibly relevant information (3). To extend the example in more detail, the query could consist of a number of recently typed words, which is then executed on an indexed collection of documents that are associated with the currently active node.

![Figure 8. Schematic of a possible scenario for the autonomous delivery of timely information relevant to the current computer based activity](image)

There are a number of other approaches for exploiting this historical record of information use in order to autonomously deliver timely and relevant information, which are out of the scope of this paper. These approaches are discussed in more detail in the related publication [1].

However, one aspect of autonomous delivery approaches, which is briefly given some consideration here, is the manner in which the computer user is alerted or presented with related information found autonomously. The categories of presentation considered in this research are as follows:

- **Passive**: In this case there is no interruption to the users current task. The information delivered is presented to the user if the user decides to check whether any information has been found related to the current context of the activity.
• **Semi-passive:** In this case the user is alerted that new information has been found, but in a way that is not intended to distract them from their current activity. (e.g. through a sound signal or an icon appearing in the computer user’s taskbar)

• **Alert:** In this case the method used to alert the user is intended to distract them from their current activity.

Depending on the user's preferences and a measure of how relevant or important the information is to the current activity, a suitable method of presentation should be adopted.

Having discussed the major components of the methodology, the next 2 sections discuss current and future work looking to validate its usefulness in practice.

### 5. Methodology Validation

2 aspects of the methodology need to be validated in order to assess its value in design practice, namely: (1) the feasibility of implementing the required software infrastructure such that all programmatic aspects of the methodology can be supported, (2) An appraisal of the value of the methodology to the individual, team and organisation. This should take into account aspects of the usability and the perceived and actual benefits of its use.

The research conducted in these areas is discussed in the following sections.

#### 5.1 Technical feasibility of implementation

A number of software components have been developed for the purposes of the research, exploiting existing technologies that enable the capture of various forms of profile element. Also, a number of other technologies have been reviewed to assess their potential role in the methodology. In summary, the following work has been carried out:

**The development of software components for:**

- The indexing, classification and search of documents stored locally (using Google and Google Desktop APIs\(^2\) [11]).
- The extraction of external content (networked or web-based) viewed through a web browser (using the IBM HTTP Web Intermediary API [9]).
- A time stamped record of computer files being edited or viewed (using Microsoft Windows audit policies)
- The monitoring of computer interface actions including text entry, and application focus (using Microsoft Windows Active Accessibility API [10])

**Feasibility studies on:**

- How the access of internal data contained within computer aided design applications such as the order and timing of feature and component creation can be achieved. The studies were based on reviews of Pro/ENGINEER, Solid Edge and SolidWorks APIs and have shown that the extraction of such data is feasible.

---

\(^2\) APIs (Application Program Interfaces) are programmable interfaces sometimes distributed or made available with software applications in order that the software can be customised, extended or controlled with additional coding.
The work has demonstrated that a number of different types of information can be successfully extracted from computer based design sessions and so there is unlikely to be any technical constraints in this regard.

However, it is anticipated that the main difficulty is likely to lie in the ability to extract data from previously unspecified sets of applications. Although there is no problem in standardising data into compatible formats, it becomes difficult to develop a generic solution that is guaranteed to be able to extract the required data from the applications that a user may happen to be using in their task which may differ considerably depending on the user and task.

5.2 Appraisal of value and benefits of the methodology

Work has been carried out with the aim of making some initial assessments about the value of using the methodology in certain scenarios. In this section, the findings from an experiment where a design activity profile was created for a mock design session are presented. The findings indicate some of the possible benefits, however, in their present form are rather subjective and are not based on data captured in an industrial setting. Plans for rolling out the testing and evaluation of the methodology in an industrial setting are discussed towards the end of this section.

Activity profile created from mock design task:

The task set for the experimental work involved the detailed and embodiment design of a small mechanical assembly. This task involved the modelling of components and assemblies in a CAD package, the requirement for the use of ISO standards material searched for and retrieved through the Internet and the sourcing of standard components, also searched for and retrieved through the Internet. This combination of modelling and information seeking activities was designed to be representative of a typical design session for this part of the design process. (Although an assumption is made that information seeking activities are primarily carried out through computer based search and retrieval methods).

The profile elements that were recorded during the design task included the URL and title of web pages retrieved, keyboard entries, application focus and the order and timing of the creation of features and components in the assembly.

Having completed the design task the data collected was compiled and analysed. A general summary of a portion of the recorded data can be seen in graphical form in Figure 9 where the shaded areas represent periods of time where the corresponding application had the focus or information had been captured.

![Figure 9. Graphical representation of application focus and data recorded through the course of the experiment](image-url)
Although further research is required to validate findings fully, the following tentative conclusions are made from the analysis.

- Superimposing a record of the timing and order of features and components created with a record of periods where information has been accessed, viewed or edited helps correlate information usage with the design of features and components within the assembly.

- Typed text in dialogue boxes and keywords used in search queries can often be used to represent the current information needs of the designer. Of the 16 text items entered throughout the course of the activity 11 queries were considered useful for forming a query describing information needs in the evaluation carried out by the author.

- An initial assessment of command usage and usage patterns in CAD packages indicates that in some scenarios command usage can be an effective prompt for providing design reference material useful to a designer and in other scenarios they can be an effective prompt for providing help in using the software application.

**Roll-out of methodology appraisal in an industrial setting**

The author is currently collaborating with the knowledge management team at Airbus UK, who are the industrial partner for this research. It is currently the intention to extend the testing and evaluation with designers within the organisation.

The author envisages that protocol analysis and survey based methods may provide the information required for an effective evaluation of the methodology. A detailed definition of the methods is currently the focus of this research, however, in summary the methods being considered are described below.

- Protocol analysis method: In this study it is anticipated that a designer will carry out a detailed or embodiment design task with the relevant software for capturing profile elements in place. In parallel, visual observations will be undertaken by the author to try and establish the underlying information needs at different stages of the design task. Having collected data from the software and from visual observations, an assessment of how effectively the profile elements can be used to provide a detailed description of the session and respond to information needs can be made.

- Trail usage and survey method: In this approach it is anticipated that the methodology will be adopted by a small project team. Qualitative and quantitative feedback on how well the systems performed can be obtained through suitable survey methods. One limitation of this approach is that a detailed history of information usage, required for the methodology to be effective, is unlikely to be in place in the current information systems infrastructure.

6. **Overall Results**

The initial results of the research carried out to date suggest that:

- In order to determine the current goal of a designer using a computer, informational elements such as the current focus of a computer application, a description of commands executed and the text typed in emails, documents and queries can be used.

- In order to provide a description of events that can be stored and reused for reference in similar future design scenarios, information elements such as the timing and order of actions in design sessions, particularly related to the access of electronic reference material is useful.
• The recording of such data using tools in the form of existing and prototypical software is feasible and should not necessarily hinder the user a great amount. Although it is acknowledged that a small proportion of time is required to facilitate the review and submission of data to nodes.

• It is suggested that issues of privacy present a possible barrier to using this extra information. (i.e. people concerned about their actions being monitored and recorded for later reuse).

7. Future work
The work carried out to develop software capabilities required by the methodology has consisted of the development of individual components carrying out the profile capture functions independently. One of the next steps in the research is to combine these independent functions so that a single interface for the review and submission of profile elements to their respective nodes can be used.

A second aspect of future work is the continuation of the evaluation and testing of the methodology discussed in the previous section.

8. Conclusions
This research has established a definition and description of data, which can be extracted but is not currently made use of, during the process of computer-based activities related to design. The research suggests that the use and storage of this data may help support the current and future information needs of designers. Perhaps the most significant benefit of profiling activities is that it is a step towards managing information about argumentation leading to decisions. Some researchers (e.g. Ullman [12]) believe this to be an important factor for the future of design information management.

9. Acknowledgements
The research carried out for this paper was funded by the Engineering and Physical Sciences Research Council (EPSRC) and Airbus UK. The authors acknowledge and thank Dr Florence Sellini and Peter Coleman from Airbus UK for their support.

References


Mr Duncan R. Campbell
Mechanical Engineering Dept.
University of Bath
Bath
BA2 7AY
England
Phone: +44 (0)1225 386131
Email: D.R.Campbell@bath.ac.uk